

Understanding Correlation Techniques for Face Recognition: From Basics to Applications

A. Alfalou¹ and C. Brosseau²

1- ISEN Brest, L@bISEN, 20 rue Cuirassé Bretagne
CS 42807, 29228 Brest Cedex 2, France

2- Université Européenne de Bretagne, Université de Brest,
Lab-STICC and Département de Physique,
CS 93837, 6 avenue Le Gorgeu, 29238 Brest Cedex 3, France

1. Introduction

This chapter covers some of the latest developments in optical correlation techniques for face recognition using the concept of spectral fusion and placing a special emphasis on its practical aspects and applications for face recognition. Optical correlation is reviewed as a method to carry out instantly a decision on the target form (form to be recognized). A range of relevant practical topics is discussed, such as JTC (Joint Transform Correlator) and the Vander-Lugt architectures. Both of them are based on the "4f" optical setup and on a comparison between the target image and a reference image. The similarity between the two images is achieved by the detection of a correlation peak. The development of suitable algorithms for optical correlation and new technologies in electro-optical interfaces has significantly improved the processing capacity of optical correlators and reduce their sizes. To overcome the limitations of a decision taken by a simple detection of a correlation peak, more complex decision algorithms are required. These algorithms necessitate the integration of new correlation filters and the realization of multiple correlations and reconfigurable multi-channel architectures.

This chapter surveys also the main correlation filters and the main multi-decision correlator architectures. It will describe the development of a multi-reconfigurable architecture based of a new version of correlation filter, i.e. the segmented filter. More specifically, we will describe an all-optical and compact correlator and a new phase optimized filter which based on four ingredients: POF, sector, composite and segmented filters. To extend the application of the algorithm and improve this system, a correlation filter is adapted in order to recognize a 3D face using multiple cameras arranged in a pickup grid. To minimize the effect of overlapping between the different spectra, the filter optimization is realized by shifting different reference spectral images. This could increase the number of references included in the same filter.

2. Background and Notations of optical correlation

The use of optical correlation methods has evolved over the past thirty years. Thus, their scope has extended to security applications, as for example tracking and/or identifying for military applications (aircraft recognition, boats recognition, etc) or civilian applications (road signs recognition, face identification: bank, metro, airport). This interest for optical correlation techniques, based on the use of two Fourier transform (FT), is essentially due to the fact that it is possible to achieve an optical Fourier Transform instantly [1] (using a simple converging lens).

To perform optical correlation techniques, two major approaches of correlators are proposed and validated in the literature: JTC (Joint Transform Correlator) [2] and Vander Lugt correlator [3]. Both approaches are based:

- 1) On an all-optical set-up, called "4f" set-up [1],
- 2) On a comparison between the target image and reference image coming from a learning data-base,
- 3) On a simple detection of a correlation peak. The latter measures the similarity degree between the target image and the reference one.

We begin this section by recalling the principles of these two approaches:

2.1 Principle of the all optical filtering 4f setup

The 4f set-up (Fig. 1) is an all optical system composed of two convergent lenses. The 2D object O (displayed in the input plane) is illuminated by a monochromatic wave [1]. A first lens performs the FT of the input object O in the image focal plane (Fourier plane), S_O . In this focal plane, a specific filter H is positioned (using optoelectronic interfaces). Next, a second convergent lens performs the inverse Fourier transform (FT^{-1}) in the output plane of the system to get the filtered image O' captured using a CCD (charge-coupled device) camera.

2.2 JTC : Joint transform correlator

Weaver and Goodman introduced the foundations of a new type of correlator for pattern recognition called JTC [2]. In their pioneering article [2], Weaver and Goodman demonstrated the possibility and the necessary conditions to achieve optically the convolution between two images displayed in the input plane of the JTC correlator: a target image (image to be recognized) and a reference image (image coming from a given data-base). Using this property and the classical "4f" set-up, Javidi and co-workers proposed [4] an optimized version of this JTC correlator. This optimized version allowing us to obtain very sharp and very intense peak correlations can increase the capacity of this type of correlators. This optimization is made possible by binarizing the joint spectrum¹ according to a well-defined threshold and using a SLM modulator. Another optimized version of the

¹ Joint spectrum: spectrum obtained after performing a FT of the input plane of the JTC correlator, containing the target image and the reference image placed at a distance d from the target [5-6].

JTC correlator was proposed by Javidi [5] by introducing a non-linearity in its Fourier plane (the JTC Fourier plane). This non-linearity in the Fourier plane increases the performances of this type of correlators.

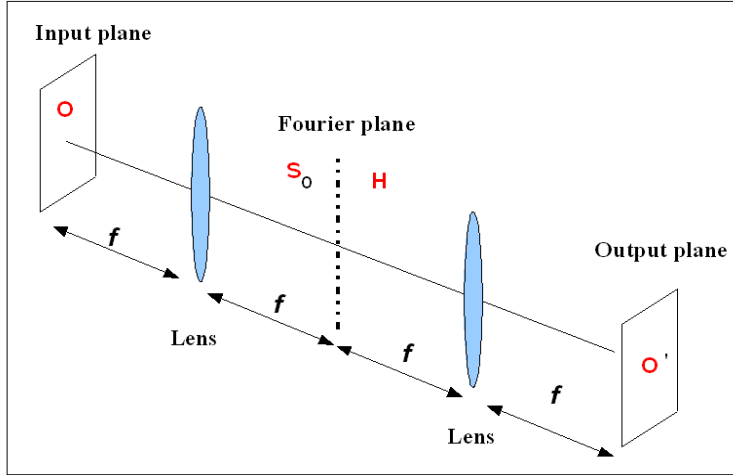


Fig. 1. All optical filtering 4f setup

The architecture of a non-linear JTC correlator (JTC-NL) [6] is now described in some detail. Other types of architecture will be listed in subsection (2.3). A synoptic diagram of the JTC set-up is presented in Fig. 2-a. Basically, it is an arrangement based upon the 4f set-up with a non-linear operation in its Fourier plane [5-6]. Its principle is to introduce, in the input plane, both the target and reference images separated by a given distance and to separate the writing and reading phases. In Ref.[6], this non-linear operation was achieved by using an Optically Addressed Spatial Light Modulator (OASLM) in the Fourier plane. A first beam coming from a laser illuminates the input plane, $I(x, y)$ Eq. (1), which contains the scene, $s(x, y)$, i.e. the target image to be recognized, and the image reference, $r(x-d, y-d)$, where d represents the distance between the target and the reference images.

$$I(x, y) = s(x, y) + r(x - d, y - d). \quad (1)$$

The joint spectrum, obtained with Fourier transformed of $I(x, y)$, is recorded by the OASLM modulator and yields $t(u, v)$ Eq. 2 (writing stage shown in Fig. 2-b).

$$t(u, v) = |S(u, v)| \times \exp[\varphi_s(u, v)] + |R(u, v)| \times \exp[\varphi_r(u, v)] \times \exp[-j(ud + vd)] \quad (2)$$

After a FT^{-1} of the joint spectrum, lighted with a second beam (reading stage shown in Fig. 2-c), the correlation between the target and the reference image on a CCD camera is obtained. The correlation plane has several peaks which correspond to the autocorrelation of

the whole scene in the center (zero order), and two correlation peaks (corresponding to the reference and target images). The location of the different correlation peaks depends on the positions of the reference and the target images within the scene.

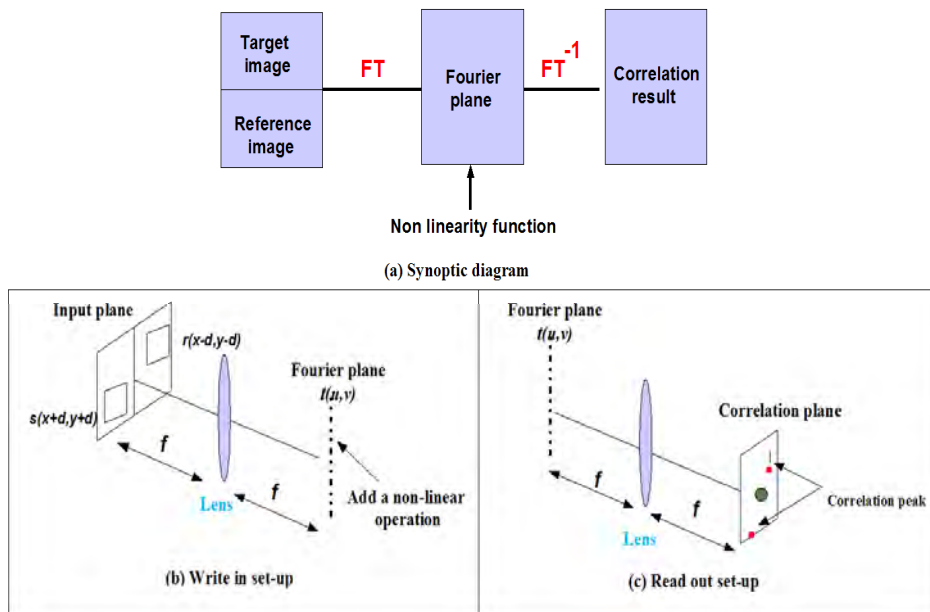


Fig. 2. Principle of the JTC correlator.

The performance of the JTC-NL correlators [5-6] proved to be larger than the linear JTC correlators [2-3]. Furthermore, by modifying the non-linearity in the Fourier plane of the JTC-NL correlator, it is possible to increase or decrease the discrimination power of the JTC-NL correlator. The interested reader may want to consult [7]. The optical implementation of the JTC correlator using an OASLAM modulator is simple to perform [6-7] and allows a high cadence (1 kHz) of processing (comparison between the target and reference images). Further adjustment in the non-linearity parameters can be easily made by changing, e.g. the voltage control of the modulator, the pulse duration, or the illumination of the modulator.

However, this correlator suffers from physical and algorithmic drawbacks. Indeed, the use of two SLM modulators in the JTC (a first one is used in the input plane to display jointly the target image and the reference one, and the other one is used in the Fourier plane to write/read the joint spectrum of JTC correlator) imposes a limited resolution depending on the technology used to manufacture these modulators. Besides the loss of information caused by these resolutions, set by the manufacturers, they also impose constraints on the choice of lens to be used for performing the FT of the input plane. The focal must verify equation (3).

$$f = \frac{N \text{Res}_{\text{input-plane}} \text{Res}_{\text{Fourier-plane}}}{\lambda}, \quad (3)$$

where $\text{Res}_{\text{input-plane}}$ denotes the resolution of the modulator used in the input plane, $\text{Res}_{\text{Fourier-plane}}$ is the resolution of the Fourier plane modulator, N is the number of pixels and λ is the wavelength. Similarly, the focal length of the second lens (to perform the second Fourier transform) is calculated by respecting the resolution of the modulator used in the Fourier plane and the resolution of the camera used to record the output plane: correlation plane.

Figure (3) shows an example of a correlation plane using a JTC-NL correlator [8]. In this plane, three peaks are visible: a central peak characterizing the zero-th order of the JTC correlator; the other two peaks are the two correlation peaks which show similar correlations between the target and reference images. The position of the two correlation peaks depends of the position of the target image relatively to the reference image placed both in the input plane.

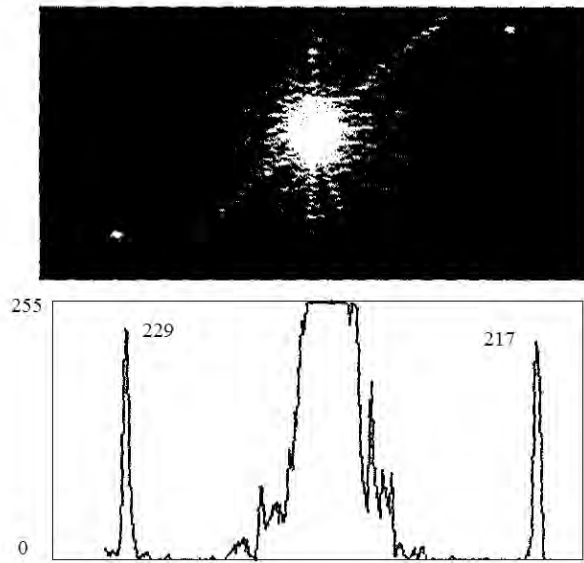


Fig. 3. Correlation plane obtained with an all Optical NL-JTC correlator [8].

2.3 The correlator JTC applied to facial recognition

To have a general overview on this type of correlator, we test by numerical simulations, some JTC architectures proposed in the literature. We examine the performance of this type of correlator vis-à-vis the face recognition application that interests us in this chapter. For this purpose, we begin by recalling the principles of several architectures implementing the

JTC correlator. All these architectures use the 4f set-up and differ from the treatment applied on the joint spectrum of the JTC correlator:

- *Classical JTC [2]: CL-JTC*

This classical architecture consists to record the intensity of the joint spectrum given by equation (2). The result is as follows

$$\begin{aligned} |t(u, v)|^2 &= |S(u, v)|^2 + |R(u, v)|^2 \\ &+ \left(|S(u, v)| \times \exp[\varphi_s(u, v)] \right) \times \left(|R(u, v)| \times \exp[-\varphi_r(u, v) + j(ud + vd)] \right) \\ &+ \left(|S(u, v)| \times \exp[-\varphi_s(u, v)] \right) \times \left(|R(u, v)| \times \exp[\varphi_r(u, v) - j(ud + vd)] \right) \end{aligned} \quad (4)$$

To obtain the correlation plane using the CL-JTC, we need to perform a FT^{-1} of the recorded intensity (Eq. 4) to get the three peaks in the correlation plane: one representing the zero-th order resulting from the FT^{-1} of the first line of Eq. 4, and two correlation peaks corresponding to the FT^{-1} of the second and third lines in Eq. 4. Studying these two correlation peaks determine the degree of resemblance between the target and the reference faces. However, this classical correlator has two major drawbacks rendering the decision of this type of correlator not always reliable:

1. It presents a very large zero-th order and very intense peak compared to the correlation peaks,
2. It is characterized by two low-intensity and large correlation peaks.

- *Binary JTC [4]: B-JTC*

Once the intensity of the joint spectrum (Eq. 4) is recorded, the binary JTC consists to binarize this intensity (Eq. 4) using the method detailed in equation (5). Then a FT^{-1} is performed to obtain the correlation plane.

$$SJ_Bin(u, v) = \begin{cases} +1 & \text{if } |t(u, v)|^2 \succ S \\ -1 & \text{if } |t(u, v)|^2 \prec S \end{cases} \quad (5)$$

where SJ_Bin is the binarized joint spectrum and S is a threshold that corresponds to the medium of the intensity joint spectrum values [4]. This correlator provides better performance than the classical JTC. It is more discriminating, has very sharp and intense correlation peaks and a very sharp zero-th order [4]. However, it is too sensitive especially to deformations of the target image relatively to the reference image (rotation, scale), which appear often in face recognition applications.

- *Nonlinear JTC [5]: NL-JTC*

To control the sensitivity of the JTC correlator, the authors of Ref [5] presented a non-linear version of this correlator. It consists in introducing a non-linearity function “ $g(E)$ ” in the joint spectrum intensity (Eq. 4). The non-linearity function $g(E)$ is given by the following equation:

$$g(E(u, v)) = E^K(u, v) \quad \text{with} \quad \begin{cases} E(u, v) = |t(u, v)|^2 \\ K \text{ is a constant value} \end{cases}, \quad (6)$$

where K represents the degree of freedom applied to our correlator. If $K=0$, we obtain the binary JTC correlate. Here we set $K=0.5$ since this value leads to a good compromise between the power discrimination and the sensitivity [7-8].

- *JTC order without zeros [9]: NZ-JTC*

As was mentioned previously, the JTC correlator has a high zero-th order which is detrimental to obtain a good and reliable decision, especially for a small distance length between the object to recognize and the reference image. Reducing this distance is required for decreasing the size of the input plane that has a limited space-bandwidth product (SBWP). Indeed, the size of the input plane depends on the size of the scene (which contains the object to be recognized), the reference image and the distance between them. To solve this problem a first technique is proposed consisting in hiding the zero-th order by multiplying the correlation plane with a binary mask given by equation (7)

$$M(x, y) = \begin{cases} +0 & \text{if } x \text{ and } y \in [-c, +c] \\ 1 & \text{if otherwise} \end{cases}, \quad (7)$$

where c is the width value of the desired mask. However, this mask cannot entirely solve the problem. Depending on the value of c , the correlation peaks can be filtered out. To overcome correctly the zero-th order problem, another technique was proposed and validated in the literature [9]. It consists in eliminating the first two terms of the joint spectrum (Eq. 4), i.e. $(|S(u, v)|^2 + |R(u, v)|^2)$, which introduce this zero-th order.

This approach was chosen by us. For that purpose, we first record separately the spectra intensities of the two images presented in the input plane: the face to be recognized and the reference. Then, we subtract the two intensity values of the equation (4) to get

$$|t(u, v)|^2 = (|S(u, v)| \times \exp[\varphi_s(u, v)]) \times (|R(u, v)| \times \exp[-\varphi_r(u, v) + j(ud + vd)]) + (|S(u, v)| \times \exp[-\varphi_s(u, v)]) \times (|R(u, v)| \times \exp[\varphi_r(u, v) - j(ud + vd)]) \quad (8)$$

This equation contains only terms corresponding to the two correlation peaks. Then, by performing a FT^{-1} of this quantity only two correlation peaks are visible.

- *Fringe-Adjusted Joint Transform Correlation: FA-JTC [10]*

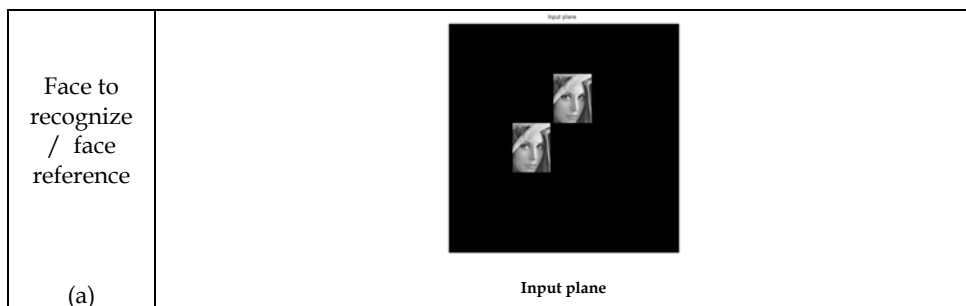
To optimize this JTC correlator and to render it more robust to noise, the authors of [10] suggested and validated a new approach of JTC, called Fringe-Adjusted Joint Transform Correlation FA-JTC. Basically, it consists in multiplying the intensity of the JTC joint spectrum (Eq 4) with the fringe-adjusted filter (Eq. 9)

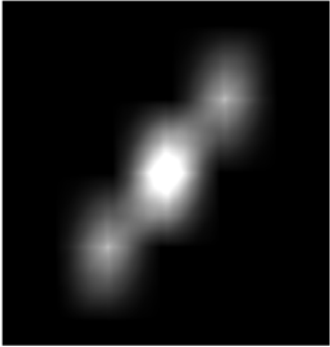
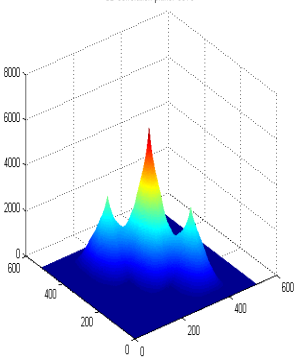

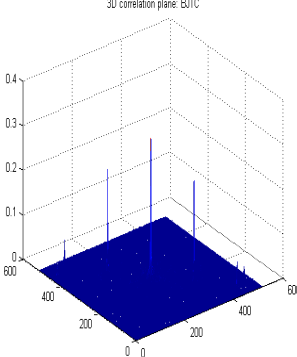
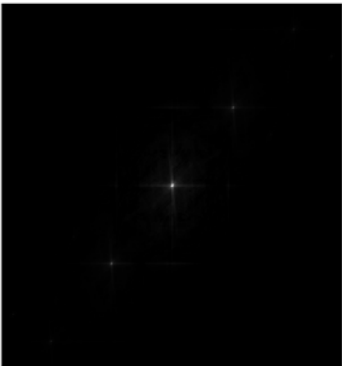
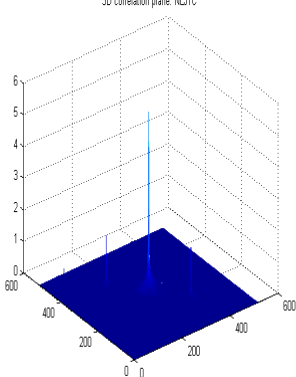
$$H(u, v) = \frac{G(u, v)}{N(u, v) + R(u, v)} , \quad (9)$$

where $G(u, v)$ denotes a function to obtain an optical gain larger than 1 and $N(u, v)$ is a function used to reduce the noise effect in the correlation peak and/or to delete the band limit the signal. A comparison of the FA-JTC method with the classical JTC and binary JTC shows that they are successful to increase the decision performance of the JTC correlator [10].

- *Comparison and conclusion*

In this section, we give an overview of the performance of several JTC correlators vis-à-vis the face recognition issue. The example chosen, Table (I-a), is to recognize the face of Lena with a reference face-image Lena. This study was conducted in a controlled environment: without noise.



<div>Classical-JTC</div> <div>(b)</div>	<div><div>correlation plane: CJTC</div></div> <div><div>3D correlation plane: CJTC</div></div> <div><div>Correlation plane: Classical JTC</div></div> <div><div>3D correlation plane: Classical JTC</div></div>
<div>B-JTC</div> <div>(c)</div>	<div><div>correlation plane: BJTC</div></div> <div><div>3D correlation plane: BJTC</div></div> <div><div>Correlation plane: B-JTC</div></div> <div><div>3D correlation plane: B-JTC</div></div>
<div>NL-JTC</div> <div>(d)</div>	<div><div>correlation plane: NLJTC</div></div> <div><div>3D correlation plane: NLJTC</div></div> <div><div>Correlation plane: NL-JTC</div></div> <div><div>3D correlation plane: NL-JTC</div></div>

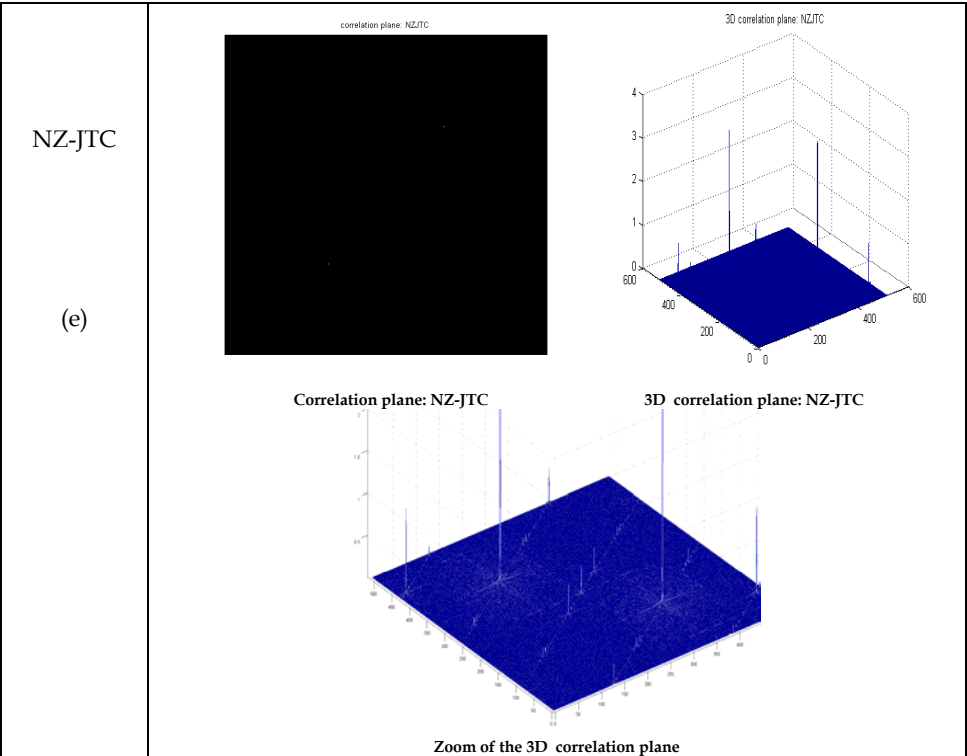


Table 1. 3D representation of the various correlation planes obtained with several approaches of the JTC correlator.

Table (1) presents information of the JTC correlator performances. The binary JTC correlator has very intense and sharp correlation peaks (Table (1-c)) compared to the values obtained with the classical JTC (Table 1-b). This results in a reliable and discriminating decision. The B-JTC correlators have good discriminating features, however they are too sensitive with respect to noise like for e.g. the deformation of the target face relative to the reference one, i.e. a small change between the reference face and the face to be recognized yields a significant decrease of the intensity of the correlation peaks. Thus, it is necessary to increase the number of reference images in the database to cover all possibilities that the target image can have in a given scene.

One solution to overcome this problem is realized by introducing a non-linearity in the Fourier plane. This allows us to control the performance of the JTC correlator (Table 1-d) and to increase or decrease its discriminating characteristic. By changing the non-linearity degree of the JTC correlator, different applications, with or without noise, can be considered. Suppression of the zero-th order can make easier the decision and increase the performance (Table 1-e).

Despite improvements found in the literature, the JTC architecture has still a serious drawback: it requires a large input *SBWP* [11]. The very nature of the JTC requires that the target and the reference (or references) images are placed side by side in the input plane (Figure 4). Assuming that N_s is the pixel size of the scene (that contains the target image with size= N_o) and N_r is the pixel size of the reference image, the pixel size of the input plane (Figure 4) N_e can be written as

$$N_e = N_s + 2 \times N_r = P \times N_r, \quad (10)$$

where P denotes the number of reference images to include in the input entry [11]. The use of many references in the input plane is necessary to overcome the problem of making a decision based on a single comparison between the target image with a single reference image. Thus, a decrease of the *SBWP_{input}* appears because of non-active zones to avoid overlapping of the correlation planes in the output plane of the JTC correlator [11]; the target and the reference images must be placed in the input plane at a distance d . Consequently, one can write $SBWP_{input} = N_e^2$, indicating that *SBWP_{input}* increases when the size of the object is increased. For example if we want to recognize a face that has a size equal to (64×64) pixels, it is necessary to use an input plane with a size equal to (320×320) pixels. This plane has an unused area equals to $8 \times (64 \times 64)$ (with reference to Figure (4)) [11].

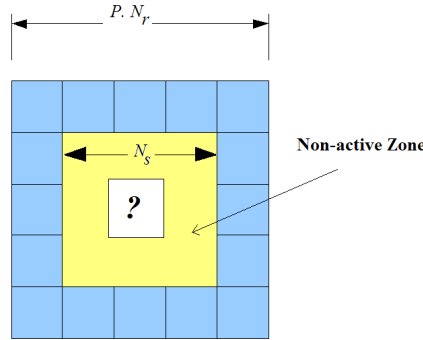


Fig. 4. Input plane of a JTC correlator using multiple references.

The implementation of an all-optical NL_JTC correlator is quite simple. It has a good compromise between discrimination and robustness; this compromise is mainly due to the use of non-linearity in the Fourier plane introduced via an optically addressed modulator (OASLM). However, the number of references that can be treated in a parallel manner is limited by the structure of the input plane (*SBWP_{input}*), i.e. the target and references must be positioned in the input plane. In addition, the use of several references in the input plane can lead to saturation in the joint spectrum, when increasing the number of references [7,11].

To overcome this problem the Vander Lugt Correlator (VLC)-based technique can be used. This technique does not require a large input plane because only the scene is presented.

Thus, all the correlation features are introduced in the Fourier plane. This technique is now described.

2.4 Vander Lugt Correlator (VLC) and mono-channel approach

The synoptic diagram of the VLC is presented in Figure 5. Basically, this technique is based on the multiplication of the spectrum, S_O , of the target image O by a correlation filter, H , made from a reference image. The input plane is illuminated with a linearly polarized incident parallel beam. Located in the rear focal plane of a Fourier lens, a SLM is used in order to display the chosen correlation filter in the Fourier plane of the optical set-up. Many approaches for designing this filter can be found in the literature according to the specific object that needs to be recognized [1,3][13-19]. A second FT is then performed with a lens in a CCD Camera (correlation plane). This results in a more or less intense central correlation peak depending on the degree of similarity between the target object and the image reference.

Since the publication of the first optical correlator VLC [3], much research has been done to render it more discriminating and robust. The 1980s and 1990s have seen a large number of correlation filters to improve the Classical Matched filter (CMF) [3]. These improvements were intended to include physical constraints of the SLM modulators which are required to display filters in the optical set-up. A few words about the notation: $I(x,y)$ and $R_i(x,y)$ are the (2D) target image (to recognize) and the 2D reference image number #i, respectively. In like fashion, $S_I(u,v)$ and $S_R(u,v)$ denote the Fourier transform of the target image and the reference image, respectively. Subsequently, we will focus on presenting several filters which have been designed for 3D face recognition.

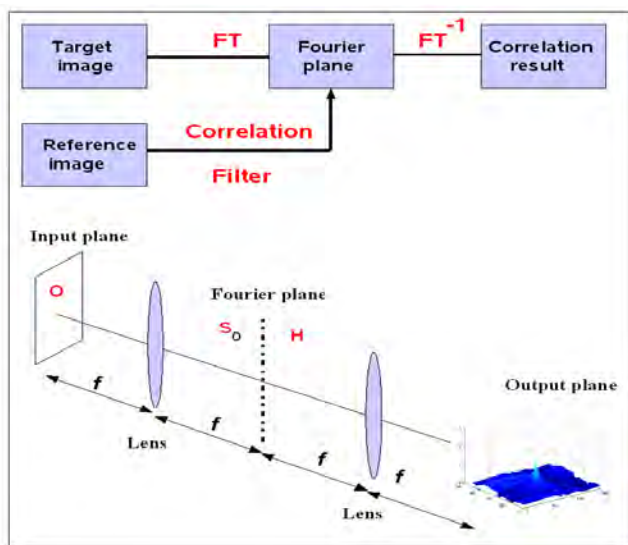


Fig. 5. Synoptic diagram of VLC.

- *Classical Matched filter: H_{CMF}*

This filter is defined as following:

$$H_{CMF}(u, v) = \frac{\alpha S_{R_i}^*(u, v)}{N(u, v)}, \quad (11)$$

where $S_{R_i}^*$ denotes the conjugate of the reference spectrum image, N is the spectral density of the background noise and α is a constant. This filter is very robust, but presents a low discriminating power [1,3,13-14]. Without noise this filter can be written as: $H_{CMF} = S_{R_i}^*(u, v)$.

- *Phase Only Filter: H_{POF}*

The POF filter is an optimized version of the classical matched filter (H_{CMF}) and is defined by equation (12). This filter has a very sharp peak, is very discriminating, but is too sensitive to specify noise arising e.g. in object deformation [13,20].

$$H_{POF} = \frac{S_{R_i}^*(u, v)}{\left| S_{R_i}^*(u, v) \right|}, \quad (12)$$

where $S_{R_i}^*(u, v)$ denotes the conjugate of the reference spectrum image and $\left| S_{R_i}^*(u, v) \right|$ is the module of the reference spectrum.

- *Binary Phase Only Filter: H_{BPOF}*

This filter is a binary version of the H_{POF} filter to take into account the physical constraints imposed by using fast and binary SLM [8,21]. Several techniques have been proposed to achieve this binarization, e.g. [21]. Here, this filter is binarized according to the following equation:

$$H_{BPOF} = \begin{cases} +1 & \text{if } \text{real}(H_{POF}) \geq 0 \\ -1 & \text{if } \text{real}(H_{POF}) < 0 \end{cases}. \quad (13)$$

Different types of binarization were considered, see e.g. Ref [22] for a review.

- *Phase Only Filter with region of support: H_{ROS_POF}*

The principle of this filter is very important for the definition of our multi-correlation segmented filter. It is defined as a multiplication of the H_{POF} filter with a pass-band function P [22-23] and can be written as following:

$$H_{ROS_POF} = H_{POF} P, \quad (14)$$

where P is a pass-band function used to select the desired information in the H_{POF} filter. This H_{ROS_POF} filter permits to increase the robustness of the classical H_{POF} , and has a good discriminating power.

The following set of filters was also proposed and discussed in the literature:

- **H_{AMPOF}**: Amplitude Matched Phase only filter [24]
- **H_{CTMF}**: Complex Ternary Matched Filter [25]
- **H_{FPP}**: Fractional power filter [14]
- **H_{IF}**: Inverse filter [14]
- **H_{PCMF}**: Phase With Constrained Magnitude filter [26]
- **H_{PMF}**: Phase mostly filter [27]
- **H_{QPF}**: Quad-Phase filter [28-29]

All these filters have been proposed to optimize the performance of the decision taken with the classical correlation filter H_{CMF} . However, in every case, each filter uses a single reference. For a reliable decision, we must compare the target image with a large number of filters. These results in a huge increase of the time required to make a decision. In addition, a small part of the *SBWP* of the output plane is used (decision based on one point in the output plane: mono-correlation). To overcome these problems and improve the reliability of the decision, multi-correlation approaches constitute a useful option [8,30].

2.5 VLC and multi-correlation approaches

Multi-correlation consists in considering a larger part of the correlation plane [8,30] to cover all possible situations and have a discriminating power. Other words stated, we explore decision-making structures based on several correlations to overcome the drawbacks existing with a decision taken with a single reference. From a practical point of view, the use of several references can be realized with a temporal multiplexing of these references (Figure 6). For that purpose, we consider successively several (number of used references) correlations to make the decision. This obviously increases the computational time.

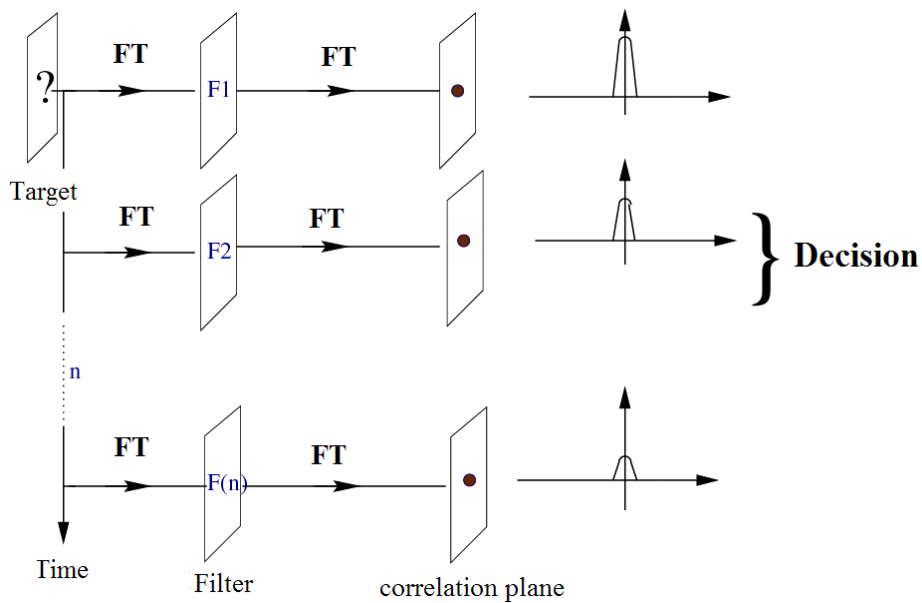


Fig. 6. Architecture using the temporal multiplexing of references.

Alfalou *et al.* [18-19] proposed and validated another solution that consists of spatially multiplexing these references together by combining their respective filters as shown in Figure (7). This solution performs simultaneously and independently many correlations. Then, the decision is taken by comparing all these correlations.

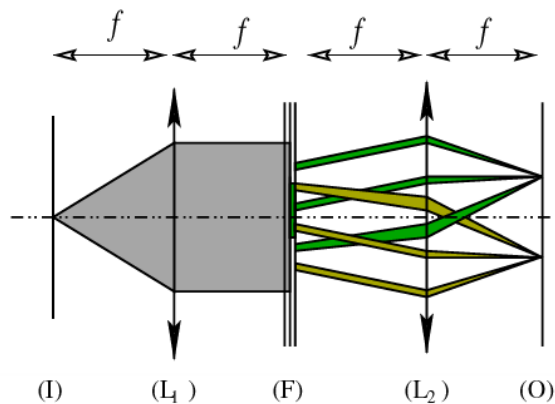


Fig. 7. Architecture using spatial multiplexing: principle of a correlator using a segmented composite filter for the multi-correlation.

Here, the filter that we seek to build is a filter allowing us to recognize one or several objects in a scene. To build such a filter, we have chosen to work with the composite filter approach for multi-correlation [18-19].

- *Composite Filter: H_{COMP}*

This filter allows one to merge multiple versions of a reference image. This type of filter is interesting to solve the correlation filter sensitivity vis-à-vis the rotation and the scaling operations. This filter can be defined as a linear combination of different reference versions:

$$H_{COM} = \sum_i a_i R_i , \quad (15)$$

where the a_i s are coefficients to optimize a cost function defined for a desired application [16-17]. In the composite filter approach, each point of the Fourier plane is involved in the correlation of a given class according to the intensity of its spectrum at this point. However, this may be inconvenient if several classes use the same point in the Fourier plane.

- *Segmented Composite Filter: H_{seg}*

To get a reliable decision, it turned out to be necessary to perform a multi-decision and reconfigurable correlator, in order to display the different references rapidly. A first solution consists in adding a grating in the correlator in order to obtain several channels at the same time [8]. However this kind of architecture lacks programmability: this is due to the fact that the various channels (necessary to separate the several decisions) are implemented in the correlator by a fixed grating. After investigating programmability opportunities of the various correlators architectures, we opted for a phase only filter based correlator [3,20] for optical implementation simplicity. But, the programmability of such architectures turned out to be a serious issue. Consequently, the study was directed towards composite filters based architectures, where the programmability introduced into the Fourier plane is easier to implement [16-17]. The various references are merged together in the same filter and can be placed in the correlator using reconfigurable interfaces. A specific carrier is assigned to each reference for separating the correlation results.

However, this needs to cope with the issue of information encoding in the Fourier plane. To perform such encoding, various schemes of the Fourier plane were proposed. One of them is adapted to the composite filters for correlation and can be carried out using the concept of segmented composite filter [18-19]. This architecture is illustrated in Figure 7. It deals with a standard optical correlator using a spatial light modulator (SLM) with binary phase-like filters in the Fourier plane.

To improve the robustness of the H_{POF} filter, it was shown that the multiplication of the latest filter with a binary pass-band function ($P(i)=0$ $i < N$ or $P(i)=1$ if $i > N$, where N is chosen according to the specific application) increases the robustness of the POF filter against noise. To improve the pass-band function, Ding and co-workers [23] proposed an iterative method. This technique requires a large computing time and depends on the desired application. Moreover, it is based on a binary function (0 or 1) leading to a loss of Horner efficiency which can be detrimental for optical implementation. A similar technique was proposed in Refs. [18-19] to improve the pass-band function based on a non-iterative method and

optimized by using the *SBWP* in the Fourier plane. This technique was shown to give a good Horner efficiency identical to a conventional H_{POF} .

This property was used to design our multi-correlation filter optimizing the use of the *SBWP* in the Fourier plane of the VLC correlator. We begin first by multiplying each filter (manufactured from one reference image) with a given pass-band function. Note that each pass-band function is chosen to eliminate all non-important areas in each filter. In these non-used areas information coming from another filter to obtain is included. Thus, a single filter containing information on several references images is used.

The significant increase in the processing capacity is due to the parallelism offered by the segmented composite filters, ensuring that an optimal use of the *SBWP* is available in the correlator. The parallelism provided by these filters offers two important advantages:

- First, it enables to increase the rate of correlations independently from the technology used for SLMs.
- Second, it offers a more efficient decision-making processes (the decision is made by the simultaneous consideration of several peaks).

• *Advantages of the segmented composite filter:*

In addition to parallelism, this filter allows one to obtain a better optimization of the *SBWP* available than conventional filters. With the composite filter, the phenomenon of local saturation in the Fourier plane is much more awkward than that met with the segmented filter, due to the fact that the manufacture of the composite filter is based on the local addition of spectral information coming from different references. On the other hand, the nature of the segmented composite filter, whose synoptic diagram is presented Figure (8), allows us to reduce the phenomenon of local saturation in Fourier plane. This result is obtained by segmenting the Fourier plane in several zones and by allotting a class to each zone. Separation at the output plane is obtained by adding a spatial carrier to each class.

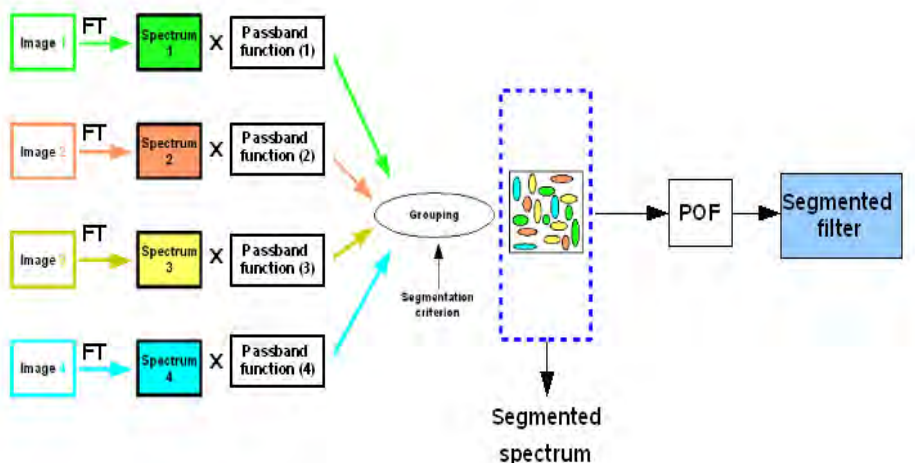


Fig. 8. Synoptic diagram of a correlator using a segmented composite filter for multi-correlation

The used criterion which was worked out in [18] is purely energetic and does not take into account the phase information. Here, the validation of such optimization of this criterion is illustrated. The decision to assign this pixel to a reference is based on the overall comparison:

$$\frac{E_{ij}^k \cos(\phi_{ij}^k)}{\sum_{i=0}^N \sum_{j=0}^N E_{ij}^k} \geq \begin{cases} \frac{E_{ij}^0 \cos(\phi_{ij}^0)}{\sum_{i=0}^N \sum_{j=0}^N E_{ij}^0} \\ \frac{E_{ij}^1 \cos(\phi_{ij}^1)}{\sum_{i=0}^N \sum_{j=0}^N E_{ij}^1} \\ \vdots \\ \frac{E_{ij}^{L-1} \cos(\phi_{ij}^{L-1})}{\sum_{i=0}^N \sum_{j=0}^N E_{ij}^{L-1}} \end{cases}, \quad (16)$$

where L is the number of references, N is the size of the filter plane, and $E_{ij}(k, l)$ is the spectral intensity of the i -th class at the pixel (k, l) . Figure (9) presents a segmented filter with a reference base made up of 4 classes (with three references per class, i.e. 9 references at all).

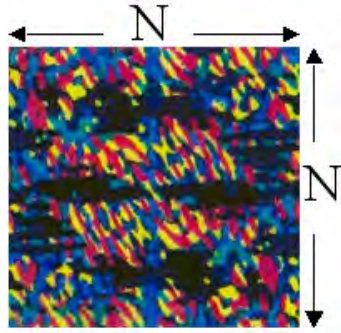


Fig. 9. Segmentation of each class in the Fourier plane

- *Criteria for assessing pattern-recognition*

Several criteria have been proposed to evaluate the correlation performance, i.e. recognition, discrimination [14, 31]. Here, two of them will be chosen for illustrative purpose. The first is the Peak-to-Correlation Energy (*PCE*) defined as the energy of the peak correlation normalized to the total energy of the correlation plane:

$$PCE = \frac{\sum_{i,j}^N E_{peak}(i, j)}{\sum_{i,j}^M E_{correlation_plane}(i, j)}, \quad (17)$$

where N denotes the size of the peak correlation spot, and M is the size of correlation plane. A second criterion is the SNR_{out} , defined as ratio of the energy of the peak correlation spot to the noise in the correlation plane:

$$SNR_{out} = \frac{\sum_{i,j}^N E_{peak}(i,j)}{\sum_{i,j}^M E_{correlation_plane}(i,j) - \sum_{i,j}^N E_{peak}(i,j)} \quad (18)$$

- *Performance of the segmented composite filter vis-à-vis the rotational invariance*

From a practical point of view, the Phase Only Segmented Composite Filter is a filter having a very good capacity of discrimination, but it is not robust. To illustrate such low robustness, we take a face onto which a rotation between 0° and 90° ($0^\circ, 10^\circ, 20^\circ, \dots, 90^\circ$) is performed. The obtained faces are then compared with a single phase segmented composite filter made from two positions of faces: 0° and with a rotation equal to 90° . Figure (10-a) confirmed the low robustness of this kind of filter, i.e. the PCE is lowered by a factor of 4. To overcome this issue, we add a reference, a face rotated by 45° in the filter. The results, incorporating the same conditions as before, are presented in figure (10-b). These results demonstrate the increase of the robustness of the segmented filter by using only three references in each class (at 0° , the second at 45° and a third to 90°).

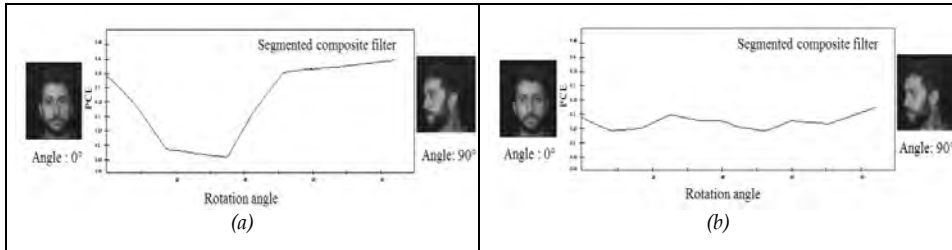


Fig. 10. Robustness of the segmented composite filter: (a) results obtained with a segmented filter used two references at 0° and 90° . (b) Results obtained with a segmented filter with three references at 0° , 45° and 90° [8]

- *Comparison between the classical composite filter and the segmented composite filter : application to face recognition*

Taking into account the limited space band-width product (SBWP), the quantity of information which can be introduced into the filter is rather limited. Consequently, the number of references which can be incorporated in the composite filter is low. Table (2) summarizes the PCE obtained numerically by introducing into the Fourier plane a single phase filter for a face recognition application. The evolution of the PCE versus the number of correlation is displayed in figure (11) (base consisted of 4 classes, with 4 references per class).

Correlation number	PCE : Composite filter	PCE : Segmented composite filter
1	0.256	0.323
2	0.163	0.309
3	0.088	0.181
4	0.066	0.141
5	0.062	0.155
6	0.033	0.138
7	0.033	0.121
8	0.029	0.121
9	0.028	0.113
10	0.029	0.099
11	0.016	0.090
12	0.012	0.090

Table 2. PCE values according various methods of calculation from filter considered

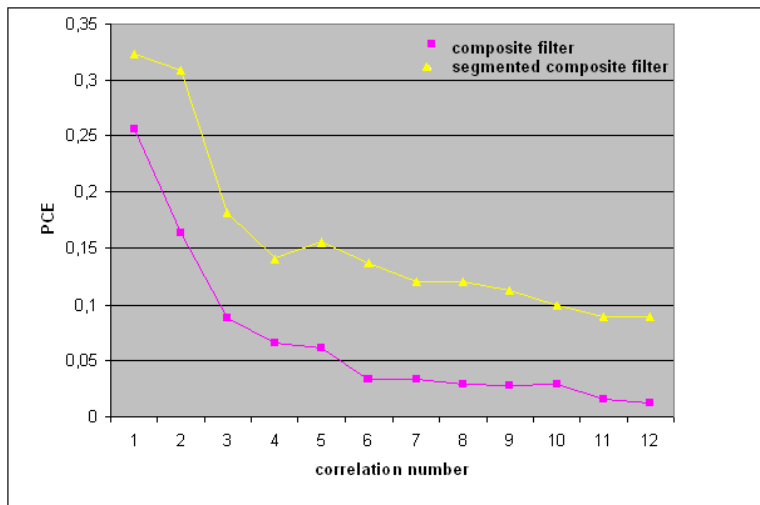


Fig. 11. PCE values obtained versus correlation number for two composite filters

In figure (11) the PCE decreases as a function of the number of correlations for different filters. It is also important to observe in figure (11), the quasi-impossibility of recognizing a face (with a classical composite filter) when more than 5 references are used. While the segmented composite filter makes it possible to incorporate 12 references the PCE loses only 1/3 of its initial value.

3. Multi-view segmented filter for multi-correlation: Application to 3D face recognition

In many practical applications, e.g. face recognition, 3D objects need to be recognized. In order to identify them, they should be first converted in 2D images. In this case, SLMs, used

to display the input plane and the filters into a correlator are 2D. The 3D-2D conversion is generally achieved with a CCD camera resulting in a loss of information. To deal with this issue, we propose to use a basic correlator (VLC) and to merge information of both object shape and object depth in its Fourier plane. Two main methods for 3D object reconstruction can be found in the literature. The first is based on holography, i.e. the PSI technique [32-33]. Even though it can provide a reliable 3D object reconstruction, it is discarded here because it needs achieving complex transmittances using several specific reference waves. Furthermore, it requires the registration of these transmittances. The second method, called integral imaging technique (II) [34] is easy to achieve since it does not require complex amplitudes and acquisition of a large amount of information. Indeed, this technique is based on the registration of multi-views of a 3D target object. These different views are called elemental 2D images. They are taken using a matrix of fixed cameras (pickup grid). Each of these cameras captures a view of the 3D object (Figure (12)). Then, the elemental 2D images are used to rebuild the 3D object by a specific process [34].

3.1. Our approach to recognize 3D object

The first step in our approach consists in decomposing the 3D object into many elemental 2D images. For this purpose, we use a pickup grid, i.e. symmetric arrangement with three rows of cameras: A first row is associated to an angle of -15° view of the object, a second one to 0° , and the third one to $+15^\circ$. In our arrangement, the cameras are placed at a distance equal to 3 cm each in length and width. To validate our approach, we used objects without noise: the subject is positioned in front of the camera placed at the center. To recognize the 3D object, we can correlate each 2D elemental image with its appropriate 2D correlation filter, thus giving an elemental decision. The final decision on the 3D object is taken by combining all these elemental decisions. However, it is a very long process that requires the *a priori* knowledge of the information on the object (especially on its orientation). To deal with this problem, we propose to merge the different elemental images together to produce a single image that contains all the necessary information on the various views of the 3D object. Next, we compare this image with only one filter. Two fusion techniques were used: the first one is to carry out the composite filter and the second one is to carry out the segmented filter. Next, we will detail the process to merge the input elemental images and the fabrication of the filter.

For the dual purpose of merging the information dealing with a 3D object and reducing it in a 2D image, we start by applying the technique used in the fabrication of composite filters. The 3D/2D image obtained is a combination of different elemental images. However, a major drawback of this technique lies in its saturation problem, especially when the number of elemental images is large, which eventually leads to a decrease of the performance of the correlator. This saturation is due to the fact that the technique we want to propose here should be optically implementable. SLMs are used to display the input images and filters that require 8-bit coding of information for the SEIKO modulator (amplitude modulator, with 256 gray levels) and 2-bit coding for the Displaytech modulator. The former modulator is used to display the target image in the input plane of our correlator. The latter modulator is used to display the various filters. To overcome the saturation, the number of merged elemental images is limited to three. A fusion technique based on the segmentation of the Fourier plane is employed [18-35].

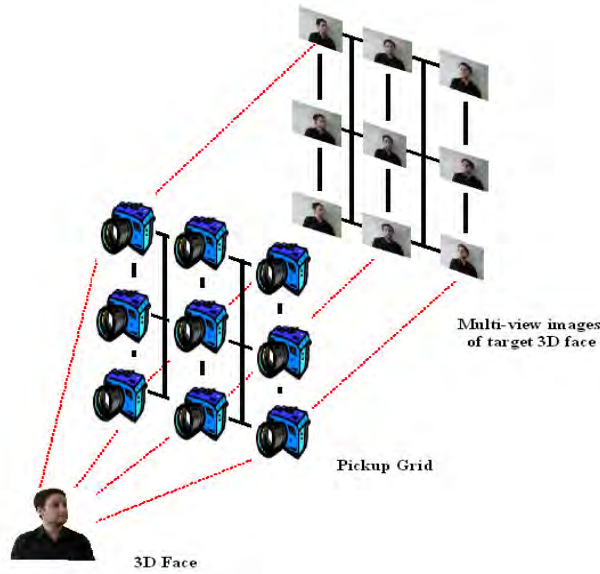


Fig. 12. Decomposition of the 3D object into several multi-view images.

3.2. Optimized multi-view segmented filter

After this short presentation of the principle of the segmented filter technique, an extension of this technique to 3D recognition is now proposed. For that purpose, we first consider a linear combination of different elemental 2D images (Figure 13-a) taken by one camera row in the pickup-grid (with reference to Figure 13-b). Next, each spectrum of merged elemental images is shifted (Figure 13-c). It was also necessary to shift the spectra to optimize the segmentation in the Fourier plane. This is required because the elemental 2D images are too close to each other. In practice, the shifting is realized optically by multiplying each merged elemental image by a specific phase term [18]. Afterwards, a segmentation to obtain the SMVSF is performed.

Particular attention was paid to find the appropriate value of the shifting parameter Δ that minimizes the overlap between the different spectra. For that purpose, we rely on previous work by Papoulis [36], who determined the necessary minimal size of a given spectrum by calculating the quantity called RMS duration

$$\Delta = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\nabla H(u, v)|^2 du dv = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x^2 + y^2) |S_I(x, y)|^2 dx dy, \quad (19)$$

where $\nabla H(u, v)$ is the gradient of the spectrum, $S_I(u, v)$ denotes the spectrum of an image I , and $1/2\pi$ is a normalization factor.

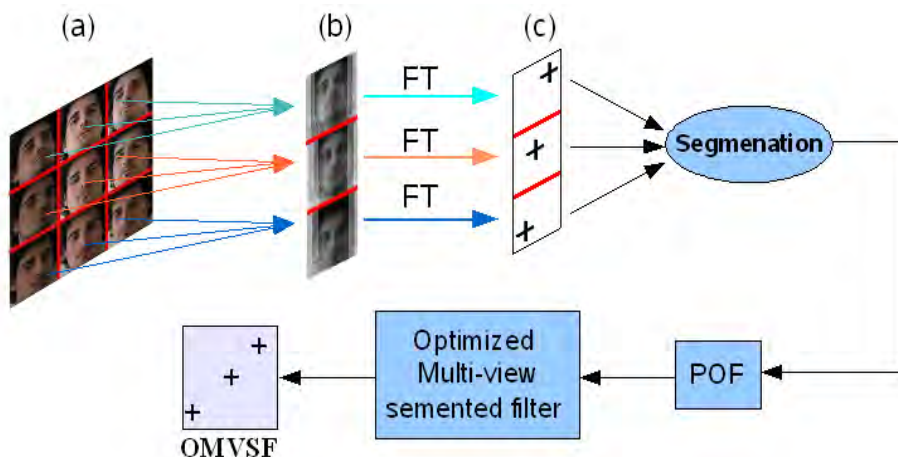


Fig. 13. Principle of the optimized multi-view segmented filter.

3.3. Proposed 3D correlator set-up

The principle of our approach is presented in Figure 14. The first operation is to add the different elemental target images. To realize this operation, we perform the same merging steps applied to fabricate our filter presented above. A multiplication with the SMVSF is followed by a FT^{-1} giving us the correlation plane. In the output plane, we get the result of simultaneous correlation of all elemental images obtained from the 3D face.

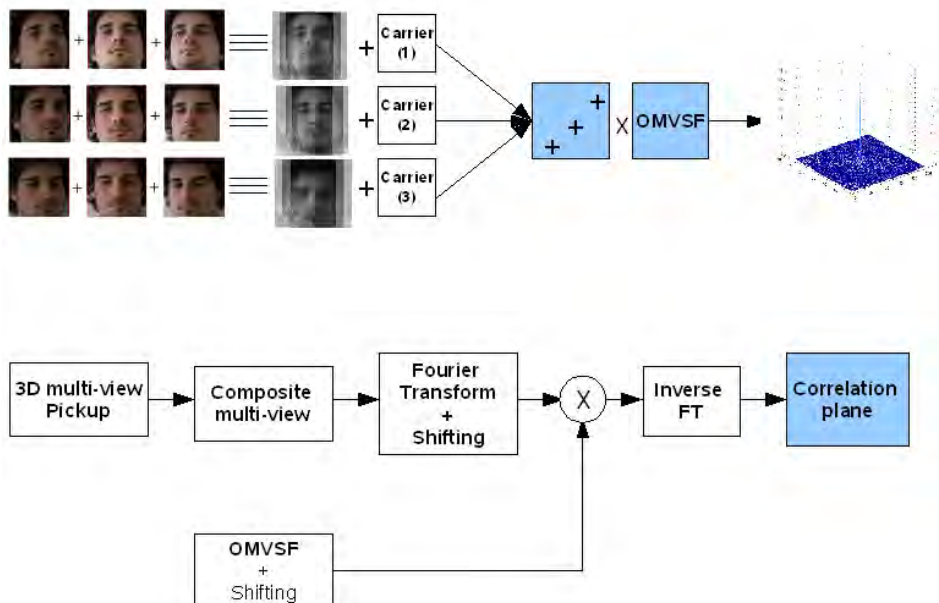


Fig. 14. Principle of the 3D Correlator set-up.

3.4. Results and Discussion

The purpose of this section is to validate the principle of our approach by numerical simulations. These simulations were conducted without noise, but taking into account the limitations imposed by using SLMs and their specific coding, i.e. 8 bit-coding of the input image and 2 bit-coding of the filter. We recall that the specific application we have in mind for these simulations is the recognition of a 3D face (with reference to Figure 15).



Fig. 15. Example of decomposition of a 3D face into multi-view 2D elemental images

If we apply our protocol for correlating the target image with a non-shifted multi-view segmented filter (Non-SMVSF), the correlation plane is presented in Figure 16(a). This figure shows the presence of a single correlation peak resulting from the autocorrelation between elemental input images with different corresponding elemental filters. It must be noted that the large width and the high level noise presented in the correlation plane (-38 dB) affects the performance of this filter. The segmentation of the different spectra is made from very similar spectra leading to strong spectral segmentation and consequently to the emergence of the isolated pixel phenomenon.

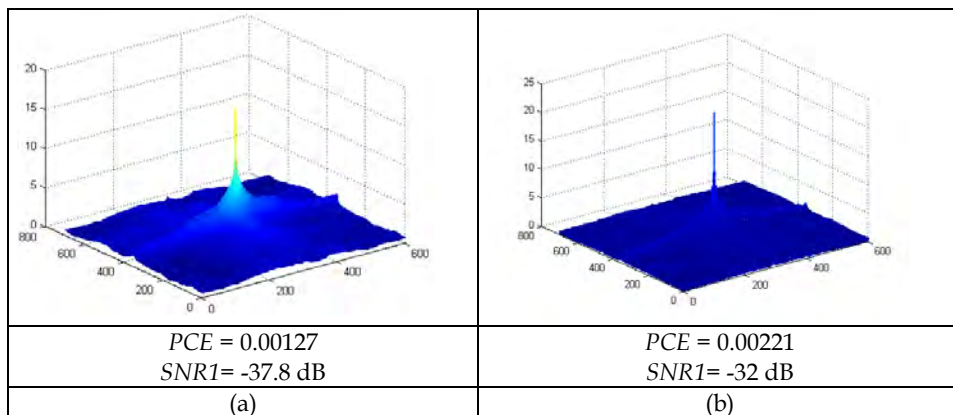


Fig. 16. Correlation plane (a) without spectra shifting, i.e. Non-SMVSF, and (b) with spectra shifting, i.e. SMVSF.

To overcome the problem of strong segmentation, we propose to optimize the use of the filter space *SBWP*. This is done by shifting the spectra to eliminate the high overlapping

areas. The shifting value was calculated using the RMS duration criterion. Figure 16-b shows the correlation plane obtained with this optimization. It is seen that the correlation peak is higher and that the noise has been significantly reduced in the correlation plane (-32 dB). From the implementation standpoint, it is necessary to binarize the filter. Figures (17-a) and (17-b) show the correlation result obtained by the binarization of two filters used in (Figure 16). The most compelling result is the significantly higher correlation peak observed in these figures.

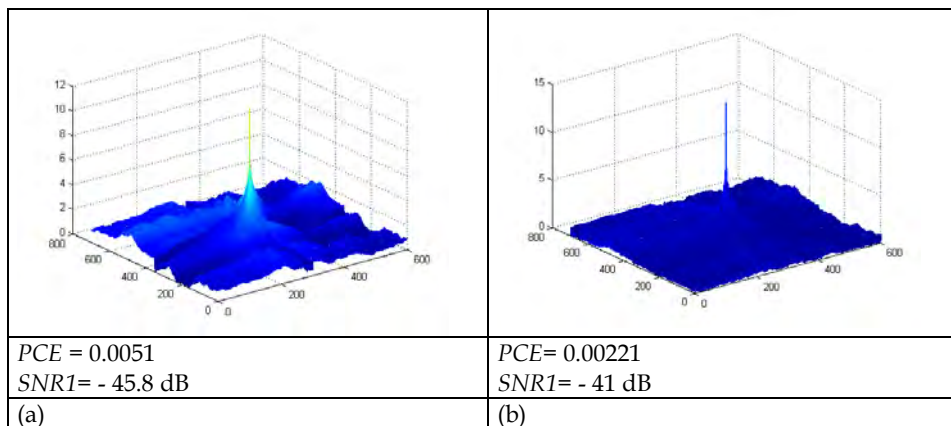


Fig. 17. Correlation plane obtained by binarizing the (a) Non-Shifting MVSF filter, and (b) the SMVSF filter, respectively.

A new approach for the purpose of recognizing 3D objects has been presented and validated. This approach is based on the decomposition of the 3D target object and the 3D reference images by using the II method. Once the elemental images are obtained, they are merged together using composite and segmented techniques. To deal with the problem of isolated pixel the different spectra were shifted. This shifting has been calculated based on the RMS duration criterion. The work presented here allows one to correlate a 3D object with a single filter, to obtain a higher and sharper correlation peak, and to reduce the noise level in the correlation plane.

4. Conclusion

In this chapter, we introduced the principle of two optical correlation approaches which can be used to recognize faces: the JTC correlator and VLC correlator. Both are based on the use of the standard optical set-up called 4f. Although the optical implementation of the JTC correlator is easier to realize than the VLC correlator, the JTC requires the use of a very large input plane in order to introduce both the face to recognize and references. Consequently, we have chosen the VLC correlator to perform our face recognition application. This choice seems appropriate, especially when a new concept of correlation filter called segmented composite filter is employed. Having detailed this filter, we presented results showing good performances of this filter applied to face recognition. Moreover, to take into account the fact that a face is a 3D object, we proposed and validated an optimization of this segmented filter suitable for 3D face recognition. For future work, it would be interesting to test the

robustness of this filter against noise and with respect to the modification of a 3D target object compared to 3D references images (rotation, scaling).

To have a broader view of face recognition problem and not to be limited only to the optical approaches, the reader may want to consult [37]. This reference presents a pure numerical approach for face recognition based on ICA (Independent Component Analysis). A comparison between these numerical and optical approaches is also provided in [37].

Finally, we point out the need of new research topics which can maximize the speed and improve the decision of the optical correlator for face recognition. Moreover, in some applications the correlator is not located in the same physical place as the target image to be recognized, thus requiring transmission and storage of the images before being processed. Therefore, it is necessary to develop appropriate techniques of image compression and encryption for recognizing a 3D face [38-39].

5. References

- [1] Goodman, J. W. (1968). *Introduction to Fourier Optics*. McGraw-Hill, New York.
- [2] Weaver C. S. & Goodman J. W. (1966). A Technique for Optically Convolution Two Functions. *Applied Optics*, 5, 1248-1249.
- [3] Vander Lugt, A. (1964). Signal detection by complex spatial filtering. *IEEE Trans. Info. Theory*, IT-10, 139-145.
- [4] Javidi, B. & Chung-Jung, K. (1988). Joint transform image correlation using a binary spatial light modulator at the Fourier plane. *Applied Optics*, 27, 4, 663-665.
- [5] Javidi, B. (1989). Nonlinear joint power spectrum based optical correlation. *Applied Optics*, 28, 2358-2367.
- [6] Guibert, L.; Keryer, G.; Serval, A.; Attia, M.; Mackenzie, H.; Pellat-Finet, P. & de Bougrenet de la Tocnaye, J. L. (1995). On-board optical joint transform correlator for real-time road sign recognition. *Optical Engineering*, 34, 101-109.
- [7] Keryer, G. (1996). *Etudes de corrélateurs optiques à corrélation jointe mono ou multicanaux : application à la reconnaissance des formes*. PhD Thesis, Université de Paris XI, Paris-France.
- [8] Alfalou, A. (1999). *Implementation of Optical Multichannel Correlators: Application to Pattern Recognition*. PhD Thesis, Université de Rennes 1-ENST Bretagne, Rennes-France.
- [9] Li, C. T.; Yin, S. & Yu, F. T. S. (1998). Nonzero-order joint transform correlator. *Optical Engineering*, 37, 58-65.
- [10] Alam, M. S. & Karim, M. A. (1993). Fringe-adjusted joint transform correlation. *Applied Optics*, 32, 4344-4350.
- [11] Keryer, G; de Bougrenet de la Tocnaye, J. L. & Al Falou, A. (1997). Performance comparison of ferroelectric liquid-crystal-technology-based coherent optical multichannel correlators. *Applied Optics*, 36, 3043-3055.
- [12] Horner, J. L. & Gianino, P.D. (1984). Phase-only matched filtering. *Applied Optics*, 23, 812-816.
- [13] Javidi, B.; Odeh, S.F. & Chen, Y. F. (1988). Rotation and scale sensitivities of the binary phase-only filter. *Applied Optics*, 27, 233-238.

- [14] Vijaya Kumar, B. V. K. & Hassebrook, L. (1990). Performance measures for correlation filters. *Applied Optics*, 29, 2997-3006.
- [15] Horner, J. L. (1992). Metrics for assessing pattern-recognition performance. *Applied Optics*, 31, 165-166.
- [16] de Bougrenet de la Tocnaye, J. L.; Quémener, E. & Pétillet, Y. (1997). Composite versus multichannel binary phase-only filtering. *Applied Optics*, 36, 6646-6653.
- [17] Kumar, B. V. K. V. (1992). Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, 31, 4773-4801.
- [18] Alfalou, A.; Keryer, G. & de Bougrenet de la Tocnaye, J. L. (1999). Optical implementation of segmented composite filtering. *Applied Optics*, 38, 6129-6136.
- [19] Alfalou, A.; Elbouz, M. & Hamam, H. (2005). Segmented phase-only filter binarized with a new error diffusion approach. *J. Opt. A: Pure Applied Optics*, 7, 183-191.
- [20] Horner J. L. & Gianino, P. D. (1984). Phase-only matched filtering. *Applied Optics*, 23, 812-816.
- [21] Horner, J. L.; Javidi, B. & Wang, J. (1992). Analysis of the binary phase-only filter. *Optics Communications*, 91, 189-192.
- [22] Kumar B. V. K. V. (1994). Partial information Filters. *Digital Signal Processing-New York*, 4, 147-153.
- [23] Ding, J.; Itoh, J. & Yatagai, T. (1995). Design of optimal phase-only filters by direct iterative search. *Optics Communications*, 118, 90-101.
- [24] Awwal, A. A. S.; Karim, M.A. & Jahan, S.R. (1990). Improved correlation discrimination using an amplitude-modulated phase-only filter. *Applied Optics*, 29, 233-236.
- [25] Dickey, F. M.; Kumar B. V. K. V.; Romero L.A. & Connelly J.M. (1990). Complex ternary matched filters yielding high signal-to-noise ratios. *Optical engineering*, 29, 9, 994-1001. ISSN 0091-3286.
- [26] Kaura, M. A. & Rhodes, W. T. (1990). Optical correlator performance using a phase-with-constrained-magnitude complex spatial filter. *Applied Optics*, 29, 2587-2593.
- [27] Richard, D. Juday, R. D. (1989). Correlation with a spatial light modulator having phase and amplitude cross coupling. *Applied Optics*, 28, 4865-4869.
- [28] Dickey, F.M. & Hansche, B. D. (1989). Quad-phase correlation filters for pattern recognition. *Applied Optics*, 28, 1611-1613.
- [29] Hansche, B. D.; Mason, J. J. & Dickey, F. M. (1989). Quad-phase-only filter implementation. *Applied Optics*, 28, 4840-4844.
- [30] Mendlovic D. & Kirysche V. (1995). Two-channel computer-generated hologram and its application for optical correlation. *Optics Communications*, 116, 322-325.
- [31] Horner, J.L. (1992). Metrics for assessing pattern-recognition performance. *Applied Optics*, 31, 165-166.
- [32] Darakis, E. & Soraghan, J. J. (2007). Reconstruction domain compression of phase-shifting digital holograms. *Applied Optics*, 46, 351-356.
- [33] Naughton, T. J.; Frauel, Y.; Javidi, B. & Tajahuerce, E. (2002). Compression of digital holograms for three-dimensional object reconstruction and recognition. *Applied Optics*, 41, 4124- 4132.
- [34] Tavakoli, B.; Daneshpanah, M.; Javidi, B. & Watson, E. (2007). Performance of 3D integral imaging with position uncertainty. *Optics Express*, 15, 11889-11902.
- [35] Farhat, M.; Alfalou, A.; Hamam, H. & Brosseau, C. (2009). Double fusion filtering based multi-view face recognition. *Optics Communications*, 282, 2136-2142.

- [36] Papoulis, A. (1962). *The Fourier Integral and its Applications*. McGrawHill, New York.
- [37] Alfalou, A.; Farhat, M. & Mansour, A. (2008). Independent Component Analysis Based Approach to Biometric Recognition. *Information and Communication Technologies: From Theory to Applications, 2008*. 7-11 April 2008 Page(s): 1-6. Digital Object Identifier 10.1109/ICTTA.2008.4530111.
- [38] Alfalou, A. & Mansour, A. (2009). A new double random phase encryption scheme to multiplex & simultaneous encode multiple images. *Applied optics*, 48, 31, 5933-5946.
- [39] Alfalou, A & Brosseau, C. (2009). Image Optical Compression and Encryption Methods. *OSA: Advances in Optics and Photonics*, 1, 589-636.